

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 06:56:21 ; Search time 149,542 Seconds
(without alignments)
1349.183 Million cell updates/sec

Title: US-09-939-537-29_COPY_1_394
Perfect score: 2029
Sequence: 1 MNRGVPRHLLVLQLALLP.....SGVLLSNIKVLPTWSTPV 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2015	99.3	458	1	CD4_HUMAN
2	1991	98.1	458	1	CD4_PANTR
3	1844	90.9	458	1	CD4_MACFU
4	1843	90.8	458	1	CD4_MACMU
5	1840	90.7	458	1	CD4_MACFA
6	1835	90.4	458	1	CD4_MACNE
7	1818	89.6	458	1	CD4_CERAE
8	1717	84.6	397	1	CD4_BRYPA
9	1707	84.1	397	1	CD4_CERTO
10	1705	84.0	397	2	009261
11	1701	83.8	397	2	009260
12	1700	83.8	397	2	009259
13	1696	83.6	397	2	009259
14	1696	83.6	397	2	009263
15	1689	83.2	397	2	009262
16	1601.5	78.9	457	2	008278
17	1575.5	77.6	457	2	008277
18	1572.5	77.5	457	2	008277
19	1183.5	58.3	455	1	CD4_SAISC
20	1182.5	58.3	455	2	0710E2
21	1146	56.5	459	1	CD4_RABIT
22	1135	55.9	463	1	CD4_CANFA
23	1119.5	55.2	457	2	0061R3
24	1099	54.1	442	2	061BKI
25	1097.5	54.1	442	2	061BKI
26	1087	53.6	474	2	P79355
27	1078.5	53.2	406	1	CD4_RAT
28	999	49.2	457	1	CD4_MOUSE
29	993	48.9	457	1	CD4_MOUSE
30	985	48.5	457	2	061396
31	954.5	47.0	433	2	055054

32	400	19.7	86	2	077596	077596 mandrill
33	400	19.7	86	2	077597	077597 mandrill
34	398	19.6	78	2	061CP8	061CP8 homo sapien
35	397	19.6	86	2	077594	077594 ceropithec
36	397	19.6	86	2	077599	077599 theropithec
37	395	19.5	86	2	077595	077595 cerocobus
38	395	19.5	86	2	071US5	071US5 cerocobus
39	392	19.3	86	2	077598	077598 papio sp. (
40	388	19.1	86	2	077601	077601 lophocobus
41	383	18.9	86	2	077600	077600 lophocobus
42	357	17.6	71	2	013969	013969 homo sapien
43	305.5	15.1	99	2	029027	029027 sus scrofa
44	280.5	13.8	99	2	029028	029028 sus scrofa
45	271.5	13.4	482	2	090WBS	090WBS anae platyr

ALIGNMENTS

RESULT 1
CD4_HUMAN STANDARD; PRT; 458 AA.
ID CD4_HUMAN
AC P01730;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
GN Name=CD4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85254948; PubMed=2990730;
RA Maddon P.J., Littman D.R., Godfrey M., Maddon D.E., Chase L., Axel R.;
RT "The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family.";
RL Cell 42:93-104(1985).
RN [2]
RP REVISION TO 26.
RX MEDLINE=89028665; PubMed=3263213; DOI=10.1016/0092-8674(88)90211-5;
RA Littman D.R., Maddon P.J., Axel R.;
RT "Corrected CD4 sequence.";
RL Cell 55:541-541(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96303695; PubMed=8723724;
RA Anasari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
MA Mailey T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase genes at human chromosome 12p13.";
RL Genome Res. 6:314-326(1996).
RN [4]
RP SEQUENCE FROM N.A. AND VARIANT TRP-265.
RX MEDLINE=91216786; PubMed=1708753; DOI=10.1016/0198-8859(91)90077-M;
RA Hodge T.W., Sasso D.R., McDougal J.S.;
RT "Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240.";
RL Hum. Immunol. 30:99-104(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,


```

FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.

Query Match 99.3%; Score 2015; DB 1; Length 458;
Best Local Similarity 99.5%; Pred. No. 2.8e-139;
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MNRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELCTASQKSIQFMKNSNOIK 60
DB 1 MNRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELCTASQKSIQFMKNSNOIK 60
OY 61 ILGNQGSFLTKGPSKLNDRADRSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
OY 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
OY 181 TWTCTVLQNKQKVEPKIDIVLAFQKASSIVYKKEGQVFPPLAFVTEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEPKIDIVLAFQKASSIVYKKEGQVFPPLAFVTEKLTGSGELMW 240
OY 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLTLPOALPOVAGSGLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLTLPOALPOVAGSGLTLA 300
OY 301 LEAKTGKHLQEVNLYVMRATQLOKNTLCEVNGPTSPKMLSLKLENKAKVSKREKPVAV 360
DB 301 LEAKTGKHLQEVNLYVMRATQLOKNTLCEVNGPTSPKMLSLKLENKAKVSKREKPVAV 360
OY 361 LNPEAGMMOCLSDSGQVLLSESNIKVLPWTSTPV 394
DB 361 LNPEAGMMOCLSDSGQVLLSESNIKVLPWTSTPV 394

RESULT 2
CD4_PANTR STANDARD; PRT; 458 AA.
AC P16004;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name:CD4;
OS Pan troglodytes (Hominidae).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; Pubmed=2107024; DOI=10.1016/0092-8674(90)90089-W;
RA Camerini D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE OF 26-424 FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=93049640; Pubmed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, M31135; AAA35407.1; -
DR EMBL, X73323; CAA51749.1; -.
DR PIR, B37322; RMC274.
DR HSSP; P01730; 1W10.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0040568; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR000973; CD4 TCAG.
DR InterPro; IPR007110; Ig_V.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_Like; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 126 125
FT DOMAIN 204 203
FT DOMAIN 317 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 42 42
FT CONFLICT 62 62
FT CONFLICT 191 191
SQ SEQUENCE 458 AA; 51057 MW; A7C3ACBA5257D3AD CRC64;

Query Match 98.1%; Score 1991; DB 1; Length 458;
Best Local Similarity 98.5%; Pred. No. 1.6e-137;
Matches 388; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 MNRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELCTASQKSIQFMKNSNOIK 60
DB 1 MNRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELCTASQKSIQFMKNSNOIK 60
OY 61 ILGNQGSFLTKGPSKLNDRADRSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
OY 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
OY 181 TWTCTVLQNKQKVEPKIDIVLAFQKASSIVYKKEGQVFPPLAFVTEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEPKIDIVLAFQKASSIVYKKEGQVFPPLAFVTEKLTGSGELMW 240
OY 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLTLPOALPOVAGSGLTLA 300

```

DB 241 QAERASSKSWITFDLKNKEVSVKRVTDPLQMGKTLPLHLTLPOLPOYAGSGNLTLLA 300
 QY 301 LEAKTGKHQEVNLYVWRAATQKNTLCEVWGPSPKLMSTLKENBAKSKREKPVWV 360
 DB 301 LEAKTGKHQEVNLYVWRAATQKNTLCEVWGPSPKLMSTLKENBAKSKREKPVWV 360
 QY 361 INPEAGMWQCLISDSGQVLTLSNLIKVLPTWSTPV 394
 DB 361 INPEAGMWQCLISDSGQVLTLSNLIKVLPTWSTPV 394

RESULT 3
 CD4_MACFU STANDARD; PRT; 458 AA.
 AC P79184;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
 GN Name=CD4;
 OS Macaca fuscata fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 OC NCBI_Taxid=9543;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hashimoto O., Tatum M.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
 CC - SUBUNIT: Associates with p56-lck (By similarity).
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

DB HSSP; D63348; BAA09672.1; -
 DR HSSP; P01730; IMBR.
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4 TCAG.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein; Palmitate; Repeat; Signal; T-cell; Transmembrane.
 KM SIGNAL 1 25 By similarity.
 FT CHAIN 26 458 T-cell surface glycoprotein CD4.
 FT DOMAIN 26 396 Extracellular (Potential).
 FT TRANSMEM 397 418 Potential.
 FT DOMAIN 419 458 Cytoplasmic (Potential).
 FT DOMAIN 26 125 Ig-like V-type.
 FT DOMAIN 126 203 Ig-like C2-type 1.
 FT DOMAIN 204 317 Ig-like C2-type 2.

FT DOMAIN 318 374 Ig-like C2-type 3.
 FT CARBOHYD 42 42 N-linked (GlcNAc...) (potential).
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 325 325 N-linked (GlcNAc...) (By similarity).
 FT DISULFID 41 109 By similarity.
 FT DISULFID 155 184 By similarity.
 FT DISULFID 328 370 By similarity.
 FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
 FT SEQUENCE 458 AA; 50828 MW; 76B3E7EF08185535 CRC64;

Query Match 90.9%; Score 1844; DB 1; Length 458;
 Best local similarity 90.4%; Pred. No. 9,5e-127; Indels 0; Gaps 0;
 Matches 356; Conservative 17; Mismatches 21;

QY 1 MNRGVPFRHLLVLTQALLPAAATQGNKVVYLGKGGDITVETCTASQKKSIOFHKNSNOIK 60
 DB 1 MNRGIPFRHLLVLTQALLPAAVTKKVVYLGKGGDITVETCTASQKKSIOFHKNSNOIK 60
 QY 61 ILGNQGSFLLTKGSPSKINDRADSRSLMDQGNFPLIKNLKIEDSDTYICVEDQKEVQL 120
 DB 61 ILGIQGSFLLTKGSPSKINDRADSRSLMDQGSFEMIINKIKIEDSDTYICEVENKKEEVEL 120
 QY 121 LVFGLTANSSTHLLQGSLLTLTSPGSPSVQCSPPGKNIQGGKTLVSQLEQDSG 180
 DB 121 LVFGLTANSSTHLLQGSLLTLTSPGSPSVQCSPPGKNIQGGKTLVSQLEQDSG 180
 QY 181 TWTCVTQONQKVEFKIDIVLAFQKASIVYKKEGQVFEFSPPLAFYVEKLTGSGELWV 240
 DB 181 TWTCVTQODQKVEFKIDIVLAFQKASIVYKKEGQVFEFSPPLAFYVEKLTGSGELWV 240
 QY 241 QAERASSKSWITFDLKNKEVSVKRVTDPLQMGKTLPLHLTLPOLPOYAGSGNLTLLA 300
 DB 241 QAERASSKSWITFDLKNKEVSVKRVTDPLQMGKTLPLHLTLPOLPOYAGSGNLTLLA 300
 QY 301 LEAKTGKHQEVNLYVWRAATQKNTLCEVWGPSPKLMSTLKENBAKSKREKPVWV 360
 DB 301 LEAKTGKHQEVNLYVWRAATQKNTLCEVWGPSPKLMSTLKENBAKSKREKPVWV 360
 QY 361 INPEAGMWQCLISDSGQVLTLSNLIKVLPTWSTPV 394
 DB 361 INPEAGMWQCLISDSGQVLTLSNLIKVLPTWSTPV 394

RESULT 4
 CD4_MACFU STANDARD; PRT; 458 AA.
 AC P16003; Q29617;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
 GN Name=CD4;
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 OC NCBI_Taxid=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W;
 RA Cemerini D., Seed B.;
 RT "A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site";
 RT Cell 60:747-754 (1990).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Thymocytes;
 RA Hashimoto O., Tatum M.;
 RT "Molecular cloning and expression of macaque CD4s";
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE OF 28-424 FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93049640; PubMed=1425921;
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
 RT "Cloning and sequences of primate CD4 molecules: diversity of the
 RT cellular receptor for simian immunodeficiency virus/human
 RT Bur. J. Immunol. 22:2973-2981 (1992).
 RN [4]
 RP SEQUENCE OF 107-192 FROM N.A.
 RX MEDLINE=98320644; PubMed=9656488;
 RA Harris B.E., Disocell T.R.;
 RT "Nuclear gene trees and the phylogenetic relationships of the
 RT mangabeys (Primates: Papionini).";
 RL Mol. Biol. Evol. 15:692-900 (1998).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL, M3134; AAA6838.1; -
 DR EMBL, D63347; BAA09671.1; -
 DR EMBL, X73326; CAA51752.1; -
 DR EMBL, AF057385; AAC25129.1; -
 DR HSP; P01730; 1MR.
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; P:coreceptor activity; ISS.
 DR GO; GO:0042289; P:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
 DR InterPro: IPR008434; CD2.
 DR InterPro: IPR000973; CD4_TcRg.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 25
 FT CHAIN 26 458
 FT DOMAIN 26 396
 FT TRANSMEM 397 418
 FT DOMAIN 419 458
 FT DOMAIN 419 458
 FT DOMAIN 126 203
 FT DOMAIN 204 317
 FT DOMAIN 318 374
 FT CARBOHYD 296 326
 FT CARBOHYD 325 325
 FT DISULFID 41 109
 FT DISULFID 155 184
 FT DISULFID 328 370
 FT LIPID 419 419
 FT LIPID 422 422
 FT CONFLICT 42 42
 FT CONFLICT 62 62
 FT CONFLICT 67 67
 FT CONFLICT 169 169
 I -> L (in Ref. 2).

FT CONFLICT 191 191 K -> N (in Ref. 3).
 FT CONFLICT 248 248 S -> P (in Ref. 2).
 FT CONFLICT 265 265 R -> Q (in Ref. 3).
 FT CONFLICT 349 349 A -> T (in Ref. 2).
 SQ SEQUENCE 458 AA; 50884 MW; 8BB80339FAFEC808 CRC64;
 Query Match 90.8%; Score 1843; DB 1; Length 458;
 Best Local Similarity 90.4%; Pred. No. 1,1e-126;
 Matches 356; Conservative 17; Mismatches 21; Indels 0; Gaps 0;
 QY 1 MNRGVPFRLHLLVLTALIPAAVQGNKVVLAGKGDVETLTASQKSIQFHWKSNQIK 60
 DB 1 MNRGIPFRLHLLVLTALIPAAVQGNKVVLAGKGDVETLTASQKSIQFHWKSNQIK 60
 QY 61 IIGNGCSFLTKGPKSLNPDARRSLMOCNPLIKNLIKIDSPYICEVDQKEVQL 120
 DB 61 IIGIQGLFTTKPKSLSDPADRSKSLMDGCSFMIKNLIKIDSPYICEVKNKEEVEL 120
 QY 121 LVFGTLNADDTLLQGSQSLTLFESPPGSSPVSQCRSPGKNIOGAGKTLVQLQLQDSG 180
 DB 121 LVFGTLNADDTLLQGSQSLTLFESPPGSSPVSQCRSPGKNIOGAGKTLVQLQLQDSG 180
 QY 181 TWCTVLAQNKVEFKIDIVLAFOKASSIVYKKEGQVFSPLAFTVEXLTGSGELMW 240
 DB 181 TWCTVLAQNKVEFKIDIVLAFOKASSIVYKKEGQVFSPLAFTVEXLTGSGELMW 240
 QY 241 QABRASSKSWITFPDLKKEVSVKRVTDPKLQNGKKPLHLTLFOALPQVAGSGNLTLA 300
 DB 241 QABRASSKSWITFPDLKKEVSVKRVTDPKLQNGKKPLHLTLFOALPQVAGSGNLTLA 300
 QY 301 LEAKTGKLGHOENVLVVWRATQLOKNUJCEVWGPTSPKMLSLKLNKAKVSKREKPVWV 360
 DB 301 LEAKTGKLGHOENVLVVWRATQLOKNUJCEVWGPTSPKMLSLKLNKAKVSKREKPVWV 360
 QY 361 LNPEAGMQLSDSGQVLLBSNINIKVPTWTPV 394
 DB 361 LNPEAGMQLSDSGQVLLBSNINIKVPTWTPV 394
 RESULT 5
 ID -MACFA STANDARD; PRT; 458 AA.
 AC P79185;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
 DE T4/Leu-3).
 DE Name=CD4;
 OS Macaca fascicularis (Cape eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymocytes;
 RA Tatemui M., Yabe M., Yamada Y.K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC EMBL: D63349; BAA09673.1; -.
DR HSSP: P01730; IMBR.
DR GO: GO:0042101; C-T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4 TCAG.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 296
FT DISULFID 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 370 370
FT LIPID 419 419
FT LIPID 422 422
SQ SEQUENCE 458 AA; 50872 MW; 9105479F85C56FF7 CRC64;
Query Match 90.7%; Score 1840; DB 1; Length 458;
Best Local Similarity 90.4%; Pred. No. 1.9e-126;
Matches 356; Conservative 16; Mismatches 22; Indels 0; Gaps 0;
1 NMRGPVFRHLVLVLOLALPAATQGNKVKLGKKDPTVELTCTASOKKSIQPHMKNSNOIK 60
1 NMRGIPFRHLVLVLOLALPAATQGNKVKLGKKDPTVELTCTASOKKSIQPHMKNSNOIK 60
61 IIGNGSFLTKGPKSLNDRADRSRLMDQGNFPLIKNLIKIDSDTYICEVEDEKERVOL 120
61 ILGIGSFLTKGPKSLNDRADRSRLMDQGNFPLIKNLIKIDSDTYICEVEDEKERVOL 120
121 LVFGILTANDTHLLOGQSITLTLESPPGSSPSVQCRSPGKNIQGAKTYSVQLELDQSG 180
121 LVFGILTANDTHLLOGQSITLTLESPPGSSPSVQCRSPGKNIQGAKTYSVQLELDQSG 180
121 LVFGILTANDTHLLOGQSITLTLESPPGSSPSVQCRSPGKNIQGAKTYSVQLELDQSG 180
181 TWTCTVLONQKVVEPKIDIVLAFQKASIVYKKKGSEVQESFPPLAFTEKLTGSGELMW 240
181 TWTCTVSOQKVVEPKIDIVLAFQKASIVYKKKGSEVQESFPPLAFTEKLTGSGELMW 240
241 QAERASSSKSMITFDLKNKEVSVKRVTDPKLQMGKQLPLHLTPALPOLVAGSGLTLTA 300
241 QAERASSSKSMITFDLKNKEVSVKRVTDPKLQMGKQLPLHLTPALPOLVAGSGLTLTA 300
301 LEAKTGKHLQEVNLYVMARFQLOKNTLGVWGPSTSPKMLSLTKENKAIVSKREKPVVW 360
301 LEAKTGKHLQEVNLYVMARFQLOKNTLGVWGPSTSPKMLSLTKENKAIVSKREKPVVW 360
361 INPEAGMOCILSDSGQVLLESNIKVLPTWSTPV 394
361 INPEAGMOCILSDSGQVLLESNIKVLPTWSTPV 394

```

```

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
ID CD4_MACNE
AC 008340; P79196;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN Name=CD4;
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_Taxid=9545;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
-----
DR EMBL: D63346; BAA09670.1; -.
DR EMBL: X73325; CAA51751.1; -.
DR HSSP: P01730; IMBR.
DR GO: GO:0042101; C-T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4 TCAG.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
SQ SEQUENCE 458 AA; 50872 MW; 9105479F85C56FF7 CRC64;

```

```

FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc...) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 328 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 57 57 D -> N (in Ref. 2).
FT CONFLICT 91 91 C -> H (in Ref. 2).
FT CONFLICT 105 105 N -> D (in Ref. 2).
FT CONFLICT 113 113 N -> D (in Ref. 2).
FT CONFLICT 302 302 D -> B (in Ref. 2).
FT CONFLICT 349 349 T -> A (in Ref. 2).
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2C8B3EB16 CRC64;

Query Match 90.4%; Score 1835; DB 1; Length 458;
Best Local Similarity 89.8%; Pred. No. 4.3e-126;
Matches 354; Conservative 19; Mismatches 21; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVLTQALPAATQGNKVLGKKGDTVELTCTPSOKKSTQFMKNSNOIK 60
DB 1 MNRGVPFRLHLVLTQALPAATQGNKVLGKKGDTVELTCTPSOKKSTQFMKNSNOIK 60
QY 61 ILGNQSFLLTKGSPSKLNDRAISRSLMDQGNFPLIKNLKIBSDTYICEVEDQKEVOL 120
DB 61 ILGNQSFLLTKGSPSKLNDRAISRSLMDQGNFPLIKNLKIBSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLQCSLTLTLESPPGSSPSVQCSPPGKNIOGSKTSLVSQLELDSDG 180
DB 121 LVFGLTANSDTHLQCSLTLTLESPPGSSPSVQCSPPGKNIOGSKTSLVSQLELDSDG 180
QY 181 TWTCTVLOKQKXVEFKIDIVLAFOKASSIVYKKEBOVFSPPLAFTEKLTGSGELMW 240
DB 181 TWTCTVLOKQKXVEFKIDIVLAFOKASSIVYKKEBOVFSPPLAFTEKLTGSGELMW 240
QY 241 QAERASSSSKWTTFDKNKEVSVKRTQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAERASSSSKWTTFDKNKEVSVKRTQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLA 300
QY 301 LEAKTGKHLQEVNLVVMRAITQLOKNTLCEVWGFSTPKMLSLKLNKKAQVSRKRPVV 360
DB 301 LEAKTGKHLQEVNLVVMRAITQLOKNTLCEVWGFSTPKMLSLKLNKKAQVSRKRPVV 360
QY 361 LNPEAGMOCILSDSGVLLSENIKTLPTWSTPV 394
DB 361 LNPEAGMOCILSDSGVLLSENIKTLPTWSTPV 394

RESULT 7
CD4_CERAE STANDARD; PRT; 458 AA.
ID CD4_CERAE STAN 458
AC Q08338; O02805; O77593; Q28217;
DC 01-FEB-1995 (Rel. 31, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_Taxid=9534;
RN [1]
RN SEQUENCE FROM N.A.
RA Hashimoto O., Tatemui M.;
RT "Molecular cloning and expression of african green monkey CD4.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;

```

```

RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus." (1992).
RL Eur. J. Immunol. 22:2973-2981 (1992).
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes."
RL J. Med. Primatol. 26:120-128 (1997).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.E., Disocell T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";
RL Mol. Biol. Evol. 15:892-900 (1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86589; BAA1332.1; -
DR EMBL; X73322; CAA51748.1; -
DR EMBL; AF001226; AAB60875.1; -
DR EMBL; AF001228; AAB60875.1; -
DR EMBL; AF057380; AAC25124.1; -
DR HSSP; P01730; 1W1Q.
DR GO; GO:0042101; C: T-cell receptor complex; ISS.
DR GO; GO:0015026; F: coreceptor activity; ISS.
DR GO; GO:0043289; F: MHC class II protein binding; ISS.
DR GO; GO:0006955; P: immune response; ISS.
DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P: T-cell differentiation; ISS.
DR GO; GO:0045058; P: T-cell selection; ISS.
DR GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR009773; CD4_TcAg.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT TRANSMEM 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 281 281 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 325 325 N-linked (GlcNAc...) (Potential).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.

```

FT DISULFID 328 370 By similarity.
FT LIPID 419 422 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 46 46 K -> N (in Ref. 2 and 3; AAB60875).
FT CONFLICT 59 59 I -> T (in Ref. 1).
FT CONFLICT 115 115 K -> B (in Ref. 3).
FT CONFLICT 165 165 G -> V (in Ref. 3; AAB60873 and 4).
FT CONFLICT 200 200 M -> V (in Ref. 2 and 3).
FT CONFLICT 227 227 F -> L (in Ref. 3; AAB60873).
FT CONFLICT 271 271 K -> H (in Ref. 3; AAB60873).
FT CONFLICT 281 281 N -> H (in Ref. 3; AAB60873).
SQ SEQUENCE 458 AA; 51158 MW; 17257 CRC64;
Query Match 89.6%; Score 1818; DB 1; Length 458;
Best Local Similarity 89.3%; Pred. No. 7, 6e-125;
Matches 352; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
QY 1 MNRGVPFRHLVLTQALALPAATQKVKVIGKKGDTVELTCTASQKKSIOFHKNSNOIK 60
DB 1 MNWGIFFRHLVLTQALALPAATQKVKVIGKKGDTVELTCTASQKKSIOFHKNSNOIK 60
QY 61 ILGNGSPFLTKGSKINDRADSRSLMDGNEPLIKLKIETSDPYICEVEDEKEVQL 120
DB 61 ILKQKSPFLTKGSKINDRADSRSLMDGNEPLIKLKIETSDPYICEVEDEKEVQL 120
QY 121 LVFGLTANSPDTHLQGSQSLTLTLESPGSSPSVQCRSPGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSPDTHLQGSQSLTLTLESPGSSPSVQCRSPGKNIQGGKTLVSQLELQDSG 180
QY 181 TWTCYVLOQKKEFKIDIVLAFQKASSTVYKKEGEVESPFLAFVTEKLTGSGELMW 240
DB 181 TWTCYVLOQKKEFKIDIVLAFQKASSTVYKKEGEVESPFLAFVTEKLTGSGELMW 240
QY 241 QAEKASSSSKMTFFDLKKEVSVKRYTOPKLOMGKLLPLHITLPOALFOYAGSGTLTA 300
DB 241 QAEKASSSSKMTFFDLKKEVSVKRYTOPKLOMGKLLPLHITLPOALFOYAGSGTLTA 300
QY 301 LEAKTGKLHQBVLVYVRAATOLQKNTLCEVWGPSTPKMLSLKLENKAKVSKREKPVV 360
DB 301 LEAKTGKLHQBVLVYVRAATOLQKNTLCEVWGPSTPKMLSLKLENKAKVSKREKPVV 360
QY 361 LNPBGMQCLISDSQVLLSENIKVLPTWPTPV 394
DB 361 LNPBGMQCLISDSQVLLSENIKVLPTWPTPV 394
RESULT 8
CD4_ERYPA
ID CD4_ERYPA STANDARD; PRT; 397 AA.
AC 008339;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)
DE (Fragment).
GN Name=CD4;
OS Erythrocytus patas (Red guenon) (Cercopithecus patas).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Erythrocebus.
OX NCBI_TaxID=9538;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus".
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).

DR EMBL: X73324; CAA51750.1; -
DR HSSP: P01730; IWO.
DR GO:GO:0042101; C: T-cell receptor complex; ISS.
DR GO:GO:0015026; F: coreceptor activity; ISS.
DR GO:GO:0042289; F: MHC class II protein binding; ISS.
DR GO:GO:0006955; P: immune response; ISS.
DR GO:GO:0045086; P: positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO:GO:0010217; P: T-cell differentiation; ISS.
DR GO:GO:0045058; P: T-cell selection; ISS.
DR GO:GO:0007169; P: transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR009733; CD4 TCAG.
DR InterPro: IPR003596; IG_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGv_1.
DR PROSITE: PS00835; IG_LIKE_1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; T-cell; Transmembrane.
FT NON_TER 1 1
FT DOMAIN 1 369
FT TRANSMEM 370 391
FT DOWAIN 392 >397
FT DOWAIN <1 98
FT DOWAIN 99 176
FT DOWAIN 177 290
FT DOWAIN 291 347
FT CARBOHYD 269 269
FT CARBOHYD 298 298
FT DISULFID 14 82
FT DISULFID 128 157
FT DISULFID 301 343
FT LIPID 392 392
FT LIPID 395 395
FT LIPID 397 397
FT NON_TER 397 397
SQ SEQUENCE 397 AA; 44081 MW; 67887397A6B7EA4F CRC64;
Query Match 84.6%; Score 1717; DB 1; Length 397;
Best Local Similarity 89.9%; Pred. No. 1.6e-117;
Matches 330; Conservative 15; Mismatches 22; Indels 0; Gaps 0;
QY 28 VILGKGDVVELTCTASQKKSIOFHKNSNOIKILGNGSPFLTKGSKINDRADSRSLW 87
DB 1 VILGKGDVVELTCTASQKKSIOFHKNSNOIKILGNGSPFLTKGSKINDRADSRSLW 87
QY 88 DQGNFPLIKLKIETSDPYICEVEDEKEVQLVRLTANSTHLLQGSQSLTLTLESP 147
DB 61 DQGNFPLIKLKIETSDPYICEVEDEKEVQLVRLTANSTHLLQGSQSLTLTLESP 120
QY 148 GSSPSVQCRSPGKNIQGGKTLVSQLELQDSGNTCTVLOQKKEFKIDIVLAFQK 207
DB 121 GSSPSVQCRSPGKNIQGGKTLVSQLELQDSGNTCTVLOQKKEFKIDIVLAFQK 180
QY 208 SSIYKKEGEVESPFLAFVTEKLTGSGELMWQAEKASSSSKMTFFDLKKEVSVKRY 267
DB 181 SSIYKKEGEVESPFLAFVTEKLTGSGELMWQAEKASSSSKMTFFDLKKEVSVKRY 240
QY 268 QDPRLQMGKLLPLHITLPOALFOYAGSGNTLALAKTGKLHQBVLVYVRAATOLQKNT 327
DB 241 QDPRLQMGKLLPLHITLPOALFOYAGSGNTLALAKTGKLHQBVLVYVRAATOLQKNT 300

QY 328 CEVWGPTSPKLMSTLKENKEAKVSRKRPVWVLPNBEAGMOCCLSDSGQVLLSNIKYL 387
 DB 301 CEVWGPTSPKLTSLKENKEATISKQAKAVWVLPNBEAGMOCCLSDSGQVLLSNIKYL 360

QY 388 PTWSTPV 394
 DB 361 PTWSTPV 367

RESULT 9
 CD4_CERTO STANDARD; PRT; 397 AA.

AC 008336;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3) (Fragment).
 GN Name=CD4;
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NC NCB1_Taxid=9531;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93049640; PubMed=1425921;
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
 RT Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus." (1992).
 RL Eur. J. Immunol. 22:2973-2981(1992).

CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.1sb-sib.ch/announce/ or send an email to license@sib-sib.ch).

DR EMBL, X73338; CA51754.1; -;
 DR EMBL, X73337; CA51753.1; -;
 DR HSBP, P01730; IWIQ.
 DR GO, GO:0042101; C:T-cell receptor complex; ISS.
 DR GO, GO:0015026; F:coreceptor activity; ISS.
 DR GO, GO:0042289; F:immune class II protein binding; ISS.
 DR GO, GO:0006955; P:immune response; ISS.
 DR GO, GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO, GO:0030217; P:T-cell differentiation; ISS.
 DR GO, GO:0045058; P:T-cell selection; ISS.
 DR GO, GO:007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4_Tcag.
 DR InterPro: IPR007110; Ig_Like.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_2.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS00835; IG_LIKE; 1.
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein; Palmitate; Repeat; T-cell; Transmembrane.
 FT NON_TER 1
 FT DOMAIN <1 369 Extracellular (Potential).
 FT TRANSMEM 370 391 Potential.

FT DOMAIN 392 >397 Cytoplasmic (Potential).
 FT DOMAIN <1 98 Ig-like V-type.
 FT DOMAIN 99 176 Ig-like C2-type 1.
 FT DOMAIN 177 230 Ig-like C2-type 2.
 FT DOMAIN 291 347 Ig-like C2-type 3.
 FT CARBOHYD 15 15 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 269 269 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 298 298 N-linked (GlcNAc...) (Potential).
 FT DISULFID 14 82 By similarity.
 FT DISULFID 128 157 By similarity.
 FT DISULFID 301 343 By similarity.
 FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
 FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
 FT VARIANT 20 20 Missing.
 FT VARIANT 43 43 T -> I.
 FT VARIANT 86 86 N -> D.
 FT VARIANT 96 96 F -> L.
 FT VARIANT 173 173 V -> M.
 FT VARIANT 316 316 R -> K.
 FT NON_TER 397 397

SO SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;

Query Match 84.1%; Score 1707; DB 1; Length 397;
 Best Local Similarity 89.6%; Pred. No. 8,6e-117;
 Matches 329; Conservative 18; Mismatches 20; Indels 0; Gaps 0;

QY 28 VYLGGKGDVVELCTPSQKKSIOFHKNKNOIKILGNQSSPLTKGSKINDRADSRSLW 87
 DB 1 VYLGGKGDVVELACNAGQKKSIOFHKNKSKQIKILGNQSSPLTKGSKINDRADSRSLW 60

QY 88 DQGNFPLITKNIKIEDSDTYICEVEDQKEVOLVGLTANSDTHLQGSITLTLESPP 147
 DB 61 DQGSFMITKNIKIEDSEYICEVENKKEVELVGLTANSDTHLQGSITLTLESPP 120

QY 148 GSSPSVQCSPPKINIQGKTLVSQLELDQSGTWTCTYLQNKVKEFKIDIVIAFOKA 207
 DB 121 GSSPSVKCSPPKINIQGRTLSVPLERQDSGTWCTVSQKVEFKIDIVIAFOKA 180

QY 208 SSTVYKKEGEVFEFSPPLAFYBKLTGSGELWMOAERASSKSWITFDLKNKEVSRYRT 267
 DB 181 SSTVYKKEGEVFEFSPPLAFYBKLTGSGELWMOAERASSKSWITFDLKNKEVSRYRT 240

QY 268 QPKLQMGKKLPLHLTLPOALPOYAGSGVTLALBAKTKLHOEVNLYVMRATOLQKNT 327
 DB 241 QPKLQMGKKLPLHLTLPOALPOYAGSGVTLALBAKTKLHOEVNLYVMRATOLQKNT 300

QY 328 CEVWGPTSPKLMSTLKENKEAKVSRKRPVWVLPNBEAGMOCCLSDSGQVLLSNIKYL 387
 DB 301 CEVWGPTSPKLTSLKENKEATISKQAKAVWVLPNBEAGMOCCLSDSGQVLLSNIKYL 360

QY 388 PTWSTPV 394
 DB 361 PTWSTPV 367

RESULT 10
 ID 009261 PRELIMINARY; PRT; 397 AA.
 AC 009261;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CD4 (Fragment).
 GN Name=CD4;
 OS Cercopithecus sabaues.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NC NCB1_Taxid=60711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98017879; PubMed=9379478;
 RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
 RT "Relation between phylogeny of African green monkey CD4 genes and
 their respective simian immunodeficiency virus genes.";
 RL J. Med. Primatol. 26:120-128(1997).
 DR EMBL: AF001225; AAB60872.1; -.
 DR HSSP: P01730; 1WIO.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0007155; P:cell adhesion; IEA.
 DR GO: GO:0006955; P:immune response; IEA.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4_TcAg.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig v.
 DR Pfam: PF05790; C2-set; 2.
 DR Pfam: PF00447; Ig 1.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 FT NON_TER 1 1
 FT NON_TER 397 397
 SQ SEQUENCE 397 AA; 43945 MW; 80C963B92A868CD3 CRC64;
 Query Match 84.0%; Score 1705; DB 2; Length 397;
 Best Local Similarity 89.9%; Pred. No. 1.2e-116;
 Matches 330; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGSPKLNDRADSRSLW 87
 DB 1 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGSKLDRIDSRKSLW 60
 QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEVOLLVPLGTANSPTHLQOGSLTLTLESP 147
 DB 61 DQCGFSMIINKLKIEDSEFTYCEVENKEEVELLVPLGTANSDTHLQOGSLTLTLESP 120
 QY 148 GSSPSVOCSPRGKNIQSGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVTLAFQKA 207
 DB 121 GSSPSVOCSPRGKNIQSGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVTLAFQKA 180
 QY 208 SSIYKKKEGEVFEFPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKRYT 267
 DB 181 SSTYKKKEGEVFEFPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKRYT 240
 QY 268 QDPKIQMGKKPLNLTLPQALPQYAGSGLTLALEAKTGKHQEVNLVWMBATQFOENLT 327
 DB 241 QDPKIQMGKKPLNLTLPQALPQYAGSGLTLALEAKTGKHQEVNLVWMBATQFOENLT 300
 QY 328 CEVWGPTSPKMLSLKLENKAKYSKREKPVVNLNPEAGMOCCLSDSGVLLSNIKVL 387
 DB 301 CEVWGPTSPKMLSLKLENKAKYSKREKPVVNLNPEAGMOCCLSDSGVLLSNIKVL 360
 QY 388 PTWSTPV 394
 DB 361 PTWSTPV 367

RESULT 11
 ID 009260 PRELIMINARY; PRT; 397 AA.
 AC 009260;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE CD4 (Fragment).
 GN Name=CD4;
 OS Cercopithecus sabaues.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecus.
 OX NCBI_TaxID=60711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98017879; PubMed=9379478;
 RA Pomegaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
 RT "Relation between phylogeny of African green monkey CD4 genes and
 their respective simian immunodeficiency virus genes.";
 RL J. Med. Primatol. 26:120-128(1997).
 DR EMBL: AF001224; AAB60871.1; -.
 DR HSSP: P01730; 1WIO.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0007155; P:cell adhesion; IEA.
 DR GO: GO:0006955; P:immune response; IEA.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4_TcAg.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig v.
 DR Pfam: PF05790; C2-set; 2.
 DR Pfam: PF00447; Ig 1.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 FT NON_TER 1 1
 FT NON_TER 397 397
 SQ SEQUENCE 397 AA; 43882 MW; 478BB27E992EE89 CRC64;
 Query Match 83.8%; Score 1701; DB 2; Length 397;
 Best Local Similarity 89.6%; Pred. No. 2.4e-116;
 Matches 329; Conservative 16; Mismatches 22; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGSPKLNDRADSRSLW 87
 DB 1 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGSKLDRIDSRKSLW 60
 QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEVOLLVPLGTANSPTHLQOGSLTLTLESP 147
 DB 61 DQCGFSMIINKLKIEDSEFTYCEVENKEEVELLVPLGTANSDTHLQOGSLTLTLESP 120
 QY 148 GSSPSVOCSPRGKNIQSGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVTLAFQKA 207
 DB 121 GSSPSVOCSPRGKNIQSGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVTLAFQKA 180
 QY 208 SSIYKKKEGEVFEFPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKRYT 267
 DB 181 SSTYKKKEGEVFEFPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKRYT 240
 QY 268 QDPKIQMGKKPLNLTLPQALPQYAGSGLTLALEAKTGKHQEVNLVWMBATQFOENLT 327
 DB 241 QDPKIQMGKKPLNLTLPQALPQYAGSGLTLALEAKTGKHQEVNLVWMBATQFOENLT 300
 QY 328 CEVWGPTSPKMLSLKLENKAKYSKREKPVVNLNPEAGMOCCLSDSGVLLSNIKVL 387
 DB 301 CEVWGPTSPKMLSLKLENKAKYSKREKPVVNLNPEAGMOCCLSDSGVLLSNIKVL 360
 QY 388 PTWSTPV 394
 DB 361 PTWSTPV 367

RESULT 12
 ID 095NE9 PRELIMINARY; PRT; 397 AA.
 AC 095NE9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE CD4 (Fragment).
 GN Name=CD4;
 OS Cercopithecus pygerythrus (Vervet monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecus.
 OX NCBI_TaxID=60710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98017879; PubMed=9379478;
 RA Pomegaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001227; AAB60874.1; -.
DR HSSP: P01730; 1W1Q.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0007155; P: cell adhesion; IEA.
DR GO: GO:0006955; P: immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcRg.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-sec; 2.
DR Pfam: PF00047; Ig; 1.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 397 AA; 43946 MW; 21C3E3082ABPBC0 CRC64;

Query Match 83.8%; Score 1700; DB 2; Length 397;
Best Local Similarity 89.6%; Pred. No. 2,8e-116;
Matches 329; Conservative 15; Mismatches 23; Indels 0; Gaps 0;

QY 28 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOTKIIGNGSFLTKGSKLNDPADSRSLW 87
DB 1 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOTKIIGNGSFLTKGSKLNDPADSRSLW 60
QY 88 DQGNFPLIINKIKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLOGQSITLTLSP 147
DB 61 DQGCFSMIINKIKIEDSEYICEVENKEEVELVFGLTANSPTHLLOGQSITLTLSP 120
QY 148 GSSPSVQCRSPRGKNIQGGKITLSVQLELDQSGTCTCYLQNKVKEFKIDIVIAFOKA 207
DB 121 GSSPSVKCRSPRGKNIQGGKITLSVQLELDQSGTCTCYLQNKVKEFKIDIVIAFOKA 180
QY 208 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 267
DB 181 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 240
QY 268 QDPKIQMGKPLNHLTPALPOYAGSGLTLALBAKTGKLHQEVNLVVMRAITQLQKNLT 327
DB 241 QDPKIQMGKPLNHLTPALPOYAGSGLTLALBAKTGKLHQEVNLVVMRAITQLQKNLT 300
QY 328 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 387
DB 301 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWSTPV 367

RESULT 13

009259 PRELIMINARY; PRT; 397 AA.
AC 009259;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus babuinus.
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_Taxid=60711;
RN [1]
RP MEDLINE=98017879; PubMed=9379478;
RX Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001223; AAB60870.1; -.
DR HSSP: P01730; 1W1Q.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0007155; P: cell adhesion; IEA.
DR GO: GO:0006955; P: immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcRg.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-sec; 2.
DR Pfam: PF00047; Ig; 1.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 397 AA; 43881 MW; 7CE39AD0F8506C81 CRC64;

Query Match 83.6%; Score 1696; DB 2; Length 397;
Best Local Similarity 89.4%; Pred. No. 5,5e-116;
Matches 328; Conservative 17; Mismatches 22; Indels 0; Gaps 0;

QY 28 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOTKIIGNGSFLTKGSKLNDPADSRSLW 87
DB 1 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOTKIIGNGSFLTKGSKLNDPADSRSLW 60
QY 88 DQGNFPLIINKIKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLOGQSITLTLSP 147
DB 61 DQGCFSMIINKIKIEDSEYICEVENKEEVELVFGLTANSPTHLLOGQSITLTLSP 120
QY 148 GSSPSVQCRSPRGKNIQGGKITLSVQLELDQSGTCTCYLQNKVKEFKIDIVIAFOKA 207
DB 121 GSSPSVKCRSPRGKNIQGGKITLSVQLELDQSGTCTCYLQNKVKEFKIDIVIAFOKA 180
QY 208 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 267
DB 181 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 240
QY 268 QDPKIQMGKPLNHLTPALPOYAGSGLTLALBAKTGKLHQEVNLVVMRAITQLQKNLT 327
DB 241 QDPKIQMGKPLNHLTPALPOYAGSGLTLALBAKTGKLHQEVNLVVMRAITQLQKNLT 300
QY 328 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 387
DB 301 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWSTPV 367

RESULT 14

009263 PRELIMINARY; PRT; 397 AA.
AC 009263;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus tantalus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_Taxid=60712;
RN [1]
RP MEDLINE=98017879; PubMed=9379478;
RX Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001221; AAB60869.1; -.
DR HSSP; P01730; 1W1Q.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR00973; CD4_TcRg.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF05790; C2-set; 2.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 397
SQ SEQUENCE 397 AA; 43994 MW; A3CD031535A51524 CRC64;

Query Match 83.6%; Score 1696; DB 2; Length 397;
Best Local Similarity 89.4%; Pred.No.5.5e-116;
Matches 328; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

```
QY 28 VVLGKGGDTVELTCTASQKKSIOFHMKNSNOIKILGNQGSFLTQGPSKLNDRADSRSLW 87
DB 1 VVLGKGGDTVELTCTASQKKSIOFHMKNSNOIKILGNQGSFLTQGPSKLNDRADSRSLW 60
QY 88 DQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVOLVFGLTANSPTHLLOGOSLTLTLESPP 147
DB 61 DQGCFSMIIKNLIKIEDSETYICEVENKKEVELLVFGLTANSPTHLLOGOSLTLTLESPP 120
QY 148 GSSPSVQCSPPKKNIOGKKTLSVSGLELODSGTCTVLOQKVEFKIDIVLAFOKA 207
DB 121 GSSPSVKCSPPKKNIOGKKTLSVSGLELODSGTCTVLOQKVEFKIDIVLAFOKA 180
QY 208 SSIIVYKKEGEQVEFSPFLAFTVEKLTGSGELMWQARASSSKSWITFDLKNKEVSKRYT 267
DB 181 SSIIVYKKEGEQVEFSPFLAFTVEKLTGSGELMWQARASSSKSWITFDLKNKEVSKRYT 240
QY 268 QDPKIQMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHGEVNLVVMRATQLOKNLT 327
DB 241 QDPKIQMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHGEVNLVVMRATQLOKNLT 300
QY 328 CEVWGTSPEKMLSLKLENKAVSKREKPVVWVNLNPEAGMOCCLSDSGQVLLSENIKYL 387
DB 301 CEVWGTSPEKMLSLKLENKAVSKREKPVVWVNLNPEAGMOCCLSDSGQVLLSENIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWPTPV 367
```

RESULT 15

009262

ID 009262 PRELIMINARY; PRT; 397 AA.

AC 009262; 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus tanzanius.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
OX NCBI_TaxId=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001221; AAB60869.1; -.
DR HSSP; P01730; 1W1Q.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR00973; CD4_TcRg.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF05790; C2-set; 2.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 397
SQ SEQUENCE 397 AA; 43994 MW; CF7F2F5D82335B0D CRC64;

Query Match 83.2%; Score 1689; DB 2; Length 397;
Best Local Similarity 89.4%; Pred.No.1.8e-115;
Matches 328; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

```
QY 28 VVLGKGGDTVELTCTASQKKSIOFHMKNSNOIKILGNQGSFLTQGPSKLNDRADSRSLW 87
DB 1 VVLGKGGDTVELTCTASQKKSIOFHMKNSNOIKILGNQGSFLTQGPSKLNDRADSRSLW 60
QY 88 DQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVOLVFGLTANSPTHLLOGOSLTLTLESPP 147
DB 61 DQGCFSMIIKNLIKIEDSETYICEVENKKEVELLVFGLTANSPTHLLOGOSLTLTLESPP 120
QY 148 GSSPSVQCSPPKKNIOGKKTLSVSGLELODSGTCTVLOQKVEFKIDIVLAFOKA 207
DB 121 GSSPSVKCSPPKKNIOGKKTLSVSGLELODSGTCTVLOQKVEFKIDIVLAFOKA 180
QY 208 SSIIVYKKEGEQVEFSPFLAFTVEKLTGSGELMWQARASSSKSWITFDLKNKEVSKRYT 267
DB 181 SSIIVYKKEGEQVEFSPFLAFTVEKLTGSGELMWQARASSSKSWITFDLKNKEVSKRYT 240
QY 268 QDPKIQMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHGEVNLVVMRATQLOKNLT 327
DB 241 QDPKIQMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHGEVNLVVMRATQLOKNLT 300
QY 328 CEVWGTSPEKMLSLKLENKAVSKREKPVVWVNLNPEAGMOCCLSDSGQVLLSENIKYL 387
DB 301 CEVWGTSPEKMLSLKLENKAVSKREKPVVWVNLNPEAGMOCCLSDSGQVLLSENIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWPTPV 367
```

Search completed: March 7, 2005, 07:19:30
Job time : 151.542 secs